Update on ProtoDUNE data studies

Ken Herner

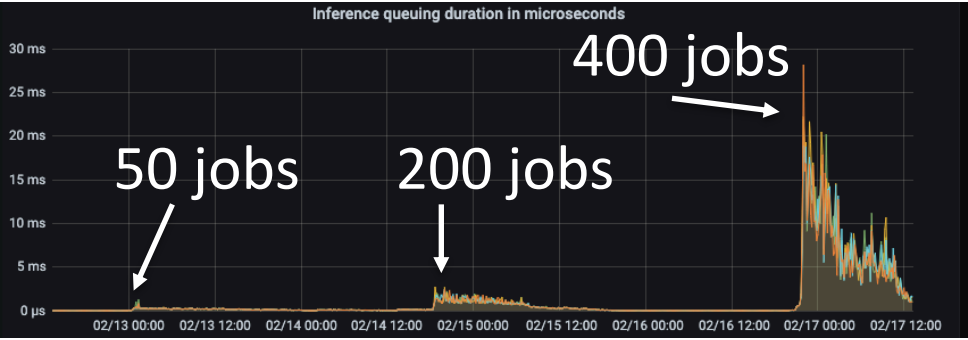
17 Feb 2021

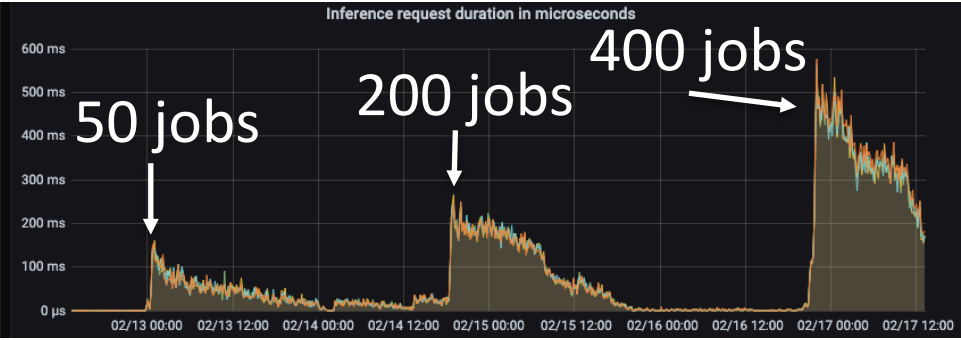
Reminder

- Similar studies as last year's paper using DUNE MC, but use detector data and try to scale up
- Using same dunetpc release as Mike before
  - Small tarball copied to each job at the beginning
- Spin up 4-GPU server in Google Cloud (Maria manages)
  - To control for network latency and data transfer effects, run only on FermiGrid (n.b. up to a factor of 2 difference in CPU speed between old and new workers)
- Found similar speedups in the fall as expected from the paper, even if perhaps not quite as good as the paper (network and disk IO contention from other jobs, etc.)

Current study

- Switched over to new funding stream, reestablished server
- Restarted with same 5 jobs to establish baseline (success)
- Scaled up to 50 simultaneous jobs, then ~200 (currently trying to do 400; expect some performance degradation at ~100 jobs per GPU)

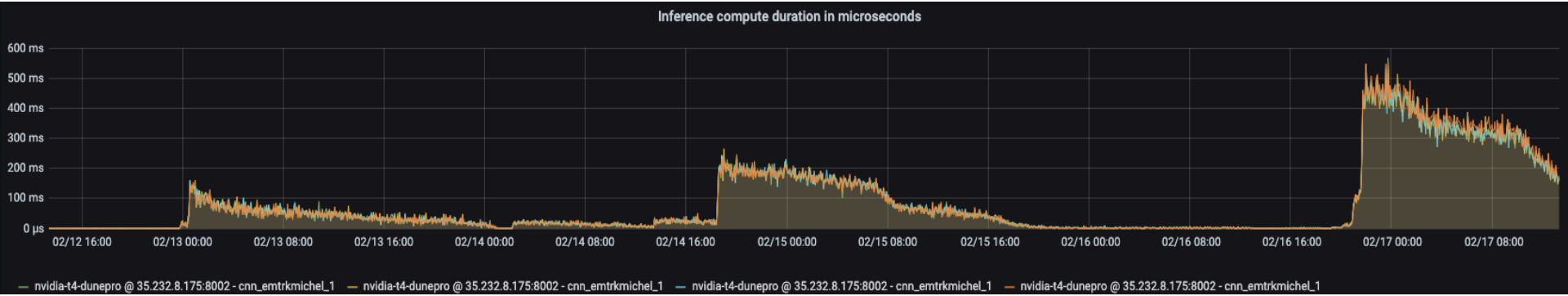




Fermilab

Current study

- Switched over to new funding stream, reestablished server
- Restarted with same 5 jobs to establish baseline (success)
- Scaled up to 50 simultaneous jobs, then ~200 (currently trying to do 400; expect some performance degradation at ~100 jobs per GPU)

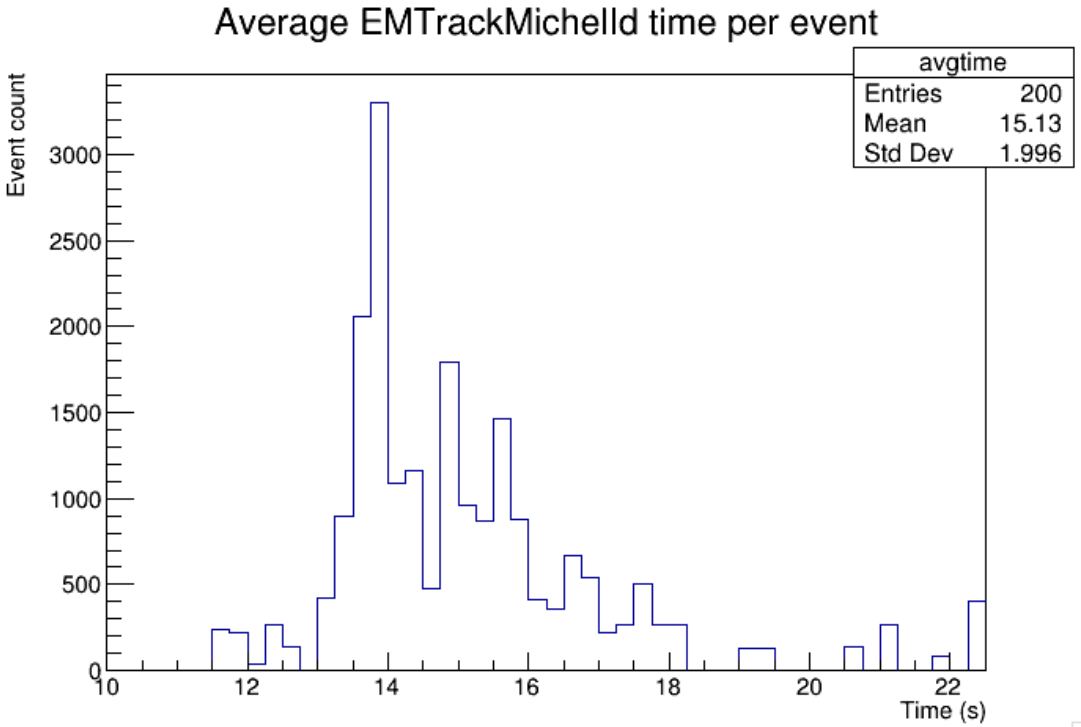


The good news: Total request time dominated by compute time

Fermilab

EmTrackMichelId Processing times

- 200 job test shown here
- Take average time in this step in each job, weight by number of events in job (not *exactly* right, but gives a rough idea)
- Performance not far off from the paper expectations!
- CPU-only average times typically closer to 200 seconds



Fermilab

Some new troubles

- Lost ~30% of jobs on the first pass in both 50 job and 200 job cases
- Largest source of error was running out of memory on the server:

Job Log

```
---- OtherArt BEGIN ---- PointIdAlgTrtis BEGIN
failed sending tRTis synchronous infer request: [inference:0 490743] INTERNAL - 2 root error(s) found.
  (0) Resource exhausted: OOM when allocating tensor with shape[14239,44,44,48] and type float on
/job:localhost/replica:0/task:0/device:GPU:0 by allocator GPU_0_bfc
    [[{{node conv2d_1/Relu-0-0-TransposeNCHWToNHWC-LayoutOptimizer}}]]
Hint: If you want to see a list of allocated tensors when OOM happens, add report_tensor_allocations_upon_oom to
RunOptions for current allocation info.
  (1) Resource exhausted: OOM when allocating tensor with shape[14239,44,44,48] and type float on
/job:localhost/replica:0/task:0/device:GPU:0 by allocator GPU_0_bfc
    [[{{node conv2d_1/Relu-0-0-TransposeNCHWToNHWC-LayoutOptimizer}}]]
Hint: If you want to see a list of allocated tensors when OOM happens, add report_tensor_allocations_upon_oom to
RunOptions for current allocation info.
    [[em_trk_none_netout/Softmax/_7]]
 Hint: If you want to see a list of allocated tensors when OOM happens, add report_tensor_allocations_upon_oom to
RunOptions for current allocation info. 0 successful operations.
```

Server Log

```
2021-02-10 02:41:26.329253: W tensorflow/core/common_runtime/bfc_allocator.cc:419] Allocator (GPU_0_bfc) ran out of
memory trying to allocate 5.67GiB (rounded to 6090129408). Current allocation summary follows.
2021-02-10 02:41:26.329302: I tensorflow/core/common_runtime/bfc_allocator.cc:869] Bin (256): Total Chunks: 5, Chunks in
use: 5. 1.2KiB allocated for chunks. 1.2KiB in use in bin. 464B client-requested in use in bin.
2021-02-10 02:41:26.329310: I tensorflow/core/common_runtime/bfc_allocator.cc:869] Bin (512): Total Chunks: 2, Chunks in
use: 2. 1.0KiB allocated for chunks. 1.0KiB in use in bin. 896B client-requested in use in bin.
2021-02-10 02:41:26.329316: I tensorflow/core/common_runtime/bfc_allocator.cc:869] Bin (1024): Total Chunks: 1, Chunks in
use: 1. 1.2KiB allocated for chunks. 1.2KiB in use in bin. 1.0KiB client-requested in use in bin.
2021-02-10 02:41:26.329325: I tensorflow/core/common_runtime/bfc_allocator.cc:869] Bin (2048): Total Chunks: 0, Chunks in
use: 0. 0B allocated for chunks. 0B in use in bin. 0B client-requested in use in bin.
2021-02-10 02:41:26.329332: I tensorflow/core/common_runtime/bfc_allocator.cc:869] Bin (4096): Total Chunks: 1, Chunks in
use: 1. 4.8KiB allocated for chunks. 4.8KiB in use in bin. 4.7KiB client-requested in use in bin.

(more such lines truncated)

2021-02-10 02:41:26.329656: I tensorflow/core/common_runtime/bfc_allocator.cc:921] Sum Total of in-use chunks: 5.73GiB
2021-02-10 02:41:26.329662: I tensorflow/core/common_runtime/bfc_allocator.cc:923] total_region_allocated_bytes_:
14220999936 memory_limit_: 14221000049 available bytes: 113 curr_region_allocation_bytes_: 68719476736
```

- Not seen at any significant level in earlier studies. *Is something different with the server, or was it just luck?*

Next Steps

- Figure out what to do with the OOM issue (bigger server?)
- Continue to scale up jobs (will also need to scale up server) at least to 1500 simultaneous